Improved maps of surface water bodies, large dams, reservoirs, and lakes in China

Xinxin Wang[1, 2], Xiangming Xiao[2, *], Yuanwei Qin[2], Jinwei Dong[3], Jihua Wu[4], Bo Li[1, 5, *]

[1] Ministry of Education Key Laboratory for Biodiversity Science and Ecological Engineering, National Observations and Research Station for Wetland Ecosystems of the Yangtze Estuary, Institute of Biodiversity Science and Institute of Eco-Chongming, School of Life Sciences, Fudan University, Shanghai 200438, China. [2] Department of Microbiology and Plant Biology, Center for Earth Observation and Modeling, University of Oklahoma, Norman, OK 73019, USA. [3] Key Laboratory of Land Surface Pattern and Simulation, Institute of Geographic Sciences and Natural Resources Research, Chinese Academy of Sciences, Beijing 100101, China. [4] State Key Laboratory of Grassland Agro-Ecosystems, and College of Ecology, Lanzhou University, Lanzhou, Gansu 730000, China. [5]Yunnan Key Laboratory of Plant Reproductive Adaptation and Evolutionary Ecology and Centre for Invasion Biology, Institute of Biodiversity, School of Ecology and Environmental Science, Yunnan University, Kunming 650504, Yunnan, China

Corresponding authors: Xiangming Xiao, xiangming.xiao@ou.edu; Bo Li, bool@fudan.edu.cn and bool@ynu.edu.cn

Abstract

Data and knowledge of surface water bodies (SWB), including large lakes and reservoirs (surface water areas > 1 km^2) are critical for the management and sustainability of water resources. However, the existing global or national dam datasets have large georeferenced coordinate offsets for many reservoirs, and some datasets have not reported reservoirs and lakes separately. In this study, we generated China's surface water bodies, Large Dams, Reservoirs, and Lakes (China-LDRL) dataset by analyzing all available Landsat imagery in 2019 (19,338 images) in Google Earth Engine and very-high spatial resolution imagery in Google Earth Pro. There were ~3.52 × 10^6 yearlong SWB polygons in China for 2019, only 0.01 × 10^6 of them (0.43%) were of large size

27 (> 1 km^2). The areas of these large SWB polygons accounted for 83.54% of the total 214.92×10^3

28 km^2 yearlong surface water area (SWA) in China. We identified 2,418 large dams, including 624

29 off-stream dams and 1,794 on-stream dams, 2,194 large reservoirs (16.35×10^3 km^2), and 3,051

30 large lakes (73.38×10^3 km^2). In general, most of the dams and reservoirs in China were distributed

31 in South China, East China, and Northeast China, whereas most of lakes were located in West

32 China, the Lower Yangtze River Basin, and Northeast China. The provision of the reliable, accurate

33 China-LDRL dataset on large reservoirs/dams and lakes will enhance our understanding of water

34 resources management and water security in China. The China-LDRL dataset is publicly available

35 at https://doi.org/10.6084/m9.figshare.16964656.v3 (Wang et al., 2022).

36 **1. Introduction**

37 Surface water bodies (SWB), including large lakes and reservoirs (surface water areas > 1

38 km^2), play an important role in the control and management of water resources (Yang and Lu, 2014,

39 2013; Feng et al., 2013, 2019). A reservoir is usually defined as artificial lake formed by

40 constructing dams across rivers (on-stream reservoir) (Thornton et al., 1996; Hayes et al., 2017)

41 or partially or completely formed by enclosed waterproof banks with concrete or clay (off-stream

42 reservoir) (Xiang et al., 2019; Thornton et al., 1996) (**Fig. S1**). Nearly 50% of the global large

43 dams were built primarily for agricultural irrigation through storing, regulating, and diverting

44 water (Mulligan et al., 2020). Additionally, they are also used for hydropower generation, human

45 and industrial uses, and flood peak attenuation (Lehner et al., 2011; Lehner and Döll, 2004; Wang

46 et al., 2021a). Large lakes have been the subject of great interest not only because of their water

47 resources but also as indicators of local climate change and anthropogenic activities (Zhang et al.,

48 2019; Ma et al., 2011; Birkett and Mason, 1995), and they could provide vital ecosystem services

49 for human being, such as alteration of river flow, supplies of irrigation water, fisheries, and

50 abundant valuable mineral deposits, and have disproportionate effects on the global carbon cycle

51 (Ran et al., 2021; Armstrong, 2010; Ma et al., 2011). Improved datasets of the numbers, sizes, and

52 spatial distributions of SWB, large dams, reservoirs, and lakes could substantially provide crucial

53 inputs for the studies of water resources, environmental health, aquatic ecosystems, and

54 agricultural sustainability (Lehner and Döll, 2004; Zhu et al., 2020).

55 **Insert Fig. 1 here**

56 China has the largest population, fastest-growing economy, increased expansion of irrigation,

57 limited water resource, dated water infrastructure, and inadequate water governance (Liu and Yang,

58 2012; Wang et al., 2020a; Tao et al., 2020). China encompasses almost 20% of the world's

59 population but contains only 7% of the world's fresh water, and as the result, it has much smaller

60 fresh water resource per capital than do most other countries (Feng et al., 2019; Dalin et al., 2014).

61 Since 1980s, China has taken diverse measures to ensure the long-term water security (Zhou et al.,

62 2020). For example, China has a remarkable increase of reservoir construction across the country

63 (Wang et al., 2021a; Zhu et al., 2020), and the total number of dams increased to ~89,700 by 2008

64 in China (Yang and Lu, 2014). The Three Gorges Reservoir, which is the world's largest

65 hydroelectric dam (Three Gorges Dam), is fully operational for flood control, power generation,

66 navigation, and water use (Wu et al., 2004; Zhang et al., 2012; Wang et al., 2013, 2020a). China

67 also has a large number of lakes with tremendously cultural and economic importance (Ma et al.,

68 2011; Zhang et al., 2019). A previous study reported that there were 2,693 large lakes (area > 1

69 km^2) in China during 2005-2006, covering 0.9% of China's land area (Ma et al., 2011). However,

70 due to intensive human activities and climate change over the last three decades, several natural

71 lakes have converted into reservoirs, dramatically accelerating shrinkage of lake areas (Yang and

72 Lu, 2014; Ma et al., 2011). Therefore, the improved datasets on the number, size, and spatial

73 distribution of large reservoirs and lakes in China is needed for assessing the impact of human

74 activities and climate change on SWB, water management, and water security in China (Zhu et al.,

75 2020; Yang and Lu, 2014).

76 Several published global dam and reservoir datasets include information from China (**Table**

77 **1**). The World Register of Dams (WRD), which was organized and released by the International

78 Commission on Large Dams (ICOLD, 2011), is the largest and widely-used dataset (Mulligan et

79 al., 2020; Paredes-Beltran et al., 2021; Wang et al., 2021a). It reports 23,841 dam entries for China,

80 however, a large proportion of those entries are not georeferenced with latitude and longitude

81 information, which limits its wide applications (Wang et al., 2021a). The GlObal GeOreferenced

82 Database of Dams (GOODD) V1 dataset reported 9,234 georeferenced dams in China (Mulligan

83 et al., 2020), however, the information (e.g. area, volume capacity) of the corresponding reservoirs

84 was not reported. The FAO's (Food and Agriculture Organization of the United Nations) global

85 information system on water resources and agricultural water management (AQUASTAT) lists

86 14,000 dams in the world, in which only part of 722 dams in China were georeferenced, and has

87 not been updated since 2015. The Global Reservoir and Dam database (GRanD), developed by the

88 Global Water System Project (GWSP), compiled the available reservoir and dam information

89 globally (Lehner et al., 2011) and has been updated for the year 2019. However, it only lists 922

90 geolocated dam entries for China. Recently, Wang et al. (2021a) released a global Georeferenced

91 global dam and reservoir (GeoDAR) dataset with 5,283 georeferenced dams in China, and the

92 reservoirs had more than 40 attributes acquired from the WRD dataset. In April 2022, the newest

93 and fully peer-reviewed version of GeoDAR is available, and this newest version had high

94 accuracy of dams in China (**Fig. S2**). There were also several published dam and reservoir maps

95 at the national scale (**Table. 1**), but these maps neither included georeferenced dams nor reported

96 reservoir attributes (e.g. reservoir area).

97 **Table 1. Information on published dam and reservoir datasets for the globe and China.**

Name	Spatial domain	Number of dams in the globe	Number of dams in China	Georeferenced dam?	Reservoir information (area …)?
WRD	Global	~ 60000	23,841	Either not georeferenced or inaccessible.	Yes, > 40 attributes
GOODD V1	Global	38667	9,231	Yes	No
FAO AQUASTAT	Global	14000	722	Partly georeferenced	Yes, reservoir capacity and area
GRanD	Global	7320	922	Yes	Yes, ~ 50 attributes
GeoDAR	Global	23680	5,283	Yes	Yes, attributes from WRD dataset
CLRM	China	/	89,700	No	Yes, reservoir capacity and area
BFNCW	China	/	98,002	No	No

98 WRD: the World Register of Dams (https://www.icold- cigb.org); GOODD: GlObal geOreferenced Database of

99   Dams (Mulligan et al., 2020); FAO AQUASTAT: The Food and Agriculture Organization of the United Nations

100   (FAO) global information system on water resources and agricultural water management

101   (http://www.fao.org/aquastat/en/databases/dams/); GRanD: the Global Reservoir and Dam database (Lehner et

102   al., 2011); GeoDAR: Georeferenced global dam and reservoir dataset (Wang et al., 2021a); CLRM: China's

103   Lakes and Reservoirs Map (Yang and Lu, 2014); BFNCW: Bulletin of First National Census for Water from

104   Ministry of Water Resources the People's Republic of China (http://www.mwr.gov.cn/2013pcgb/index.html). "/"

105   means these data were published in China, but global dam information is unavailable.

106

107   In addition to the dam and reservoir datasets, several studies have reported the spatial

108   distribution and multi-year dynamics of inland SWB (Tao et al., 2020; Ma et al., 2011; Wang et

109   al., 2020a; Feng et al., 2019) and lakes (Gao, 2015; Gao et al., 2012; Ma et al., 2011; Zhang et al.,

110   2019) in China, however, they did not explicitly explore the spatial distribution of large reservoirs

111   and lakes in China, making it impossible to assess the impact of human activities on these two

112   types of water resources. Thus, to date, the spatial distributions of SWB, large dams, reservoirs,

113   and lakes in China have not been fully investigated and documented, yet.

114   The objective of this study was to produce detailed and accurate maps of open SWB, large

115   dams, reservoirs, and lakes (surface water area > 1 km^2) in China in 2019, the latest year when this

116   study started in late 2020, and those SWB with area $\leq$ 1km^2 were excluded. First, this study used

117   time-series Landsat imagery in 2019 and Google Earth Engine (GEE) cloud computing platform

118   as well as the simple and robust surface water mapping algorithm (Zou et al., 2018, 2017; Zhou et

119 al., 2019b; Wang et al., 2020a) to generate raster maps of SWB in China at 30-m spatial resolution.

120 Second, we converted the raster map of SWB to a vector map of SWB and identified those large

121 SWB with area > 1 km^2. Third, we combined the vector maps of SWB with the historical satellite

122 images in 2019 within China in Google Earth Pro to identify dams and released China's surface

123 water bodies, large dams, reservoirs, and lakes dataset, namely, China-LDRL. Forth, we analyzed

124 the spatial distribution of SWB, large dams, reservoirs, and lakes in China. Finally, we discussed

125 the reliabilities, uncertainties, limitations, outlooks, and implications of the China-LDRL dataset

126 for the study of water security.

127

128 **2. Materials and Methods**

129 **2.1 Study area**

130   The study area covered all the provincial-level administrative divisions in China (**Fig. 1a**),

131 including 23 provinces, 2 special administrative regions (Hong Kong and Macao), 4 municipalities

132 (Beijing, Tianjin, Shanghai, and Chongqing), and 5 Autonomous Regions (Inner Mongolia,

133 Guangxi, Tibet, Ningxia, and Xinjiang). Since Macao and Hong Kong have relatively small areas

134 and are very close to Guangdong Province, we combined them as one region (Guangdong) when

135 we performed the statistical analysis in this study.

136   China has great altitude diversity as the eastern plains and southern coasts consist of lowlands

137 and foothills, the southern areas of China consist of hilly and mountainous terrains, the west and

138 north of the country are dominated by basins, plateaus, and massifs, and the southwestern China

139 contains part of the highest tablelands on earth, the Tibetan Plateau (**Fig. 1a**). Due to substantial

140 differences in latitude, longitude, and altitude, the climate of China is extremely diverse, ranging

141 from tropical in the far south to subarctic in the far north and alpine in the higher elevations of the

142 Tibetan Plateau, contributing to the much more surface water areas in Southwest and Southeast of

143 China than other regions, especially North China (Wang et al., 2020a).

144 **Insert Fig. 1 here**

145

146 **2.2 Data**

147 **2.2.1 Landsat data**

148 In this study, we used the available Landsat surface reflectance (SR) images in the GEE

149 platform, and there was a total of 19,338 images in 2019 for China, including 9,028 Landsat-7

150 ETM+ images and 10,310 Landsat-8 OLI images (~21.73 TB). The detailed information of

151 Landsat SR products is available on the GEE platform (https://developers.google.com/earth-

152 engine/datasets/catalog/landsat, last access: 18 February 2022). All these images had undergone

153 necessary pre-processing in GEE, including radiometric calibration, atmospheric correction, and

154 the removal of stripes in Landsat-7 imagery. We used the quality assurance (QA) band that was

155 generated by the CFMASK algorithm (Zhu et al., 2015) to identify bad-quality observations,

156 including clouds and cloud shadows (Murray et al., 2019; Pekel et al., 2016). We also used the

157   Shuttle Radar Topography Mission (SRTM) digital elevation model (DEM) data, the solar azimuth

158   and zenith angle data of each image, and ee.Terrain.hillShadow algorithm in GEE to identify those

159   pixels with terrain shadows (Zou et al., 2018; Wang et al., 2020a) (**Fig. 1b**), which were excluded

160   from the data analysis. Out of ~132.43 million pixels in China, approximately 98.36% had more

161   than 5 good-quality observations and 91.24% had more than 10 good-quality observations in 2019.

162   About 93.14% of the 78.9 million pixels in North China had more than 20 good-quality

163   observations due to the overlapping of Landsat images at the high latitudes and less cloud cover

164   (Zhou et al., 2019a; Wang et al., 2020b). Note that the number of Landsat-7 ETM+ images in GEE

165   may change in the future, as USGS continues to work with the International Ground stations (IGS)

166   in the world to assemble and rescue some images from individual stations. For Landsat-8 OLI

167   images, USGS does not rely on IGS for image downlink, as its data record is able to store all the

168   images and then downlink them to the Landsat archive (Wulder et al., 2016).

169   We used three spectral indices (NDVI, EVI, mNDWI) to identify SWB in this study. These

170   indices are defined as:

171     $NDVI = \frac{\rho_{nir}-\rho_{red}}{\rho_{nir}+\rho_{red}}$ (1)

172     $EVI = 2.5 \times \frac{\rho_{nir}-\rho_{red}}{\rho_{nir}+6\times\rho_{red}-7.5\times\rho_{blue}+1}$ (2)

173     $mNDWI = \frac{\rho_{green}-\rho_{swir}}{\rho_{green}+\rho_{swir}}$ (3)

174   where ρ_{blue}, ρ_{green}, ρ_{red}, ρ_{nir}, and ρ_{swir} are blue, green, red, near-infrared, and shortwave

175   infrared bands of Landsat images, respectively.

2.2.2 Dam and reservoir datasets

The GlObal GeOreferenced Database of Dams (GOODD) dataset was released in 2020 and it lists ~38,000 georeferenced dams as well as derived data on their associated catchments through one by one degree titles on the Google Earth geobrowser during 2007-2011 and the Shuttle Radar Topography Mission (SRTM) Water Body Dataset (SWBD) (Mulligan et al., 2020). It provides the raw digitized coordinates for the locations of dam walls, but it does not provide the detailed attribute data on the characteristics of each dam and reservoir (**Fig. 2a, d**). Both the large dams and medium sized dams were captured in this dataset.

The Global Reservoir and Dam (GRanD) Database v1.3 was recently updated in February 2019 by Lehner et al. (2011) (**Fig. 2b, e**). The spatial information of these dams was contributed by eleven participating institutions. Each dam was assigned to a polygon that depicted the reservoir surface, which was provided by SWBD (v1.1) and the surface water maps were produced by the Joint Research Center (JRC) of the European Commission from Landsat imagery at 30-m spatial resolution for the period 1984-2015 (Pekel et al., 2016) (v1.3). All reservoirs with a storage capacity of more than 0.1 km^3 were included in this dataset, and some smaller reservoirs were also added when their data were available.

The Georeferenced global Dam And Reservoir dataset (GeoDAR) was produced by utilizing multi-source dam and reservoir inventories (ICOLD WRD and GRanD v1.3 datasets) and the Google Maps geocoding API (Wang et al., 2021a) (**Fig. 2d, e**). The GeoDAR product includes two successive versions. GeoDAR v1.0 is essentially a georeferenced subset of ICOLD WRD, and

196 contains more than 20,000 dam entries, and each of which is indexed by an encrypted identifier

197 (ID) that is associated with a WRD record, allowing for the potential retrieval of all its 40+

198 proprietary attributes from ICOLD. GeoDAR v1.1 consists of (1) dam entries as in v1.0 except

199 those that further harmonized with GRanD for an improved inclusion of the largest dams, and (2)

200 reservoir boundaries for most of the dam entries. The GeoDAR was just updated in April 2022 and

201 is available at https://doi.org/10.5281/zenodo.6163413.

202 **Insert Fig. 2 here**

203 **2.3 Methods**

204 The workflow for producing the China-LDRL dataset included two major sections: 1)

205 generation of yearlong SWB maps in China by analyzing time-series Landsat imagery in 2019

206 with GEE platform, and 2) identification of dams and classification of yearlong SWB into lakes,

207 reservoirs, and rivers by analyzing the historical satellite images in 2019 within China in Google

208 Earth Pro. A flowchart showing the methodology of this study is illustrated in **Fig. 3**.

209 **Insert Fig. 3 here**

210 **2.3.1 Algorithm to generate annual maps of yearlong surface water bodies**

211 In this study, we combined a surface water index (mNDWI) and two greenness-based

212 vegetation indices (EVI and NDVI) to identify SWB through the algorithm of ((mNDWI > EVI or

213 mNDWI > NDVI) and EVI < 0.1) (Eq. (4)). This mNDWI/VIs algorithm can reduce the effects of

214 vegetation on identification of SWB, and has already been used to identify and map SWB at the

215   regional and national scales with high accuracy (Zou et al., 2018, 2017; Zhou et al., 2019b; Wang

216   et al., 2020a). Furthermore, this mNDWI/VIs algorithm had been compared with other surface

217   water body mapping algorithms (e.g. NDWI, mNDWI, TCW, and AWEI), and the results showed

218   that the integration of mNDWI/VIs algorithm and Landsat images could identify SWB with high

219   producer's accuracy (98.1%) and user's accuracy (91.0%) (Zhou et al., 2017).

220   Surface water body frequency (F_{SWB}) of a pixel in a year was calculated as the ratio of the

221   number of observations identified as surface water body to the number of good-quality

222   observations in a year and ranged from 0 to 1.0 (or 100%) (Zou et al., 2017), see Eq. (5). We

223   generated the F_{SWB} map of all the pixels in China for 2019 in the GEE platform (**Fig. 5a**).

224    $$SWB = \begin{cases} 1, & (\text{mNDWI>EVI or mNDWI>NDVI}) \text{ and EVI<0.1} \\ 0, & \text{Other values} \end{cases} \tag{4}$$

225    $$F_{SWB} = \frac{N_{SWB}}{N_{good}} \tag{5}$$

226   where SWB is surface water body, F_{SWB} is surface water body frequency, N_{SWB} is the number

227   of observations identified as SWB (see Eq. (4)) in 2019, N_{good} is the number of good-quality

228   observations in 2019.

229   Consistent with our previous studies (Zou et al., 2018; Wang et al., 2020a), a water pixel was

230   defined as yearlong surface water body ($F_{SWB} \geq 0.75$), seasonal surface water body ($0.05 \leq F_{SWB}$

231   < 0.75), or ephemeral surface water body ($F_{SWB} < 0.05$). We generated the seasonal and yearlong

232   SWB maps in China for 2019, respectively (**Fig. 4b, c**).

233   **Insert Fig. 4 here**

2.3.2 The procedure to identify dams, reservoirs, and lakes in Google Earth Pro

We first generated the yearlong SWB vector map in China for 2019 based on the yearlong SWB raster map, then reprojected it to the Krasovsky_1940_Albers equal-area conic projection and calculated the area (km^2) of each yearlong SWB polygon as its attribute (Python code is available in: https://drive.google.com/drive/folders/1B19VKbCIoDPmu-IcmiZcOIUF8wi1YnE?usp=sharing). When we reported large reservoirs and lakes, only those polygons with area > 1 km^2 were kept in this study (**Fig. S3**). In an effort to distinguish riverine or off-stream reservoirs from lakes, we uploaded the large SWB vector layers into Google Earth Pro, and checked whether a dam existed around each polygon through the historical satellite images in 2019 within China by visual image interpretation approach. If a dam did not exist, we classified the polygon as river or lake; if a dam does exist, we classified the polygon as on-stream reservoir (constructed on a river/stream regardless of impoundment) or off-stream reservoir (formed by partial or complete embankment around an off-stream lake) (**Fig. S1**). Simultaneously, the corresponding dam would be classified as on-stream dam or off-stream dam. Finally, the SWB polygons were classified into lakes, reservoirs, and rivers, and the dams/reservoirs were classified into on-stream and off-stream dams/reservoirs (**Fig. 3**). This work was carried out and completed by the lead author (Dr. Wang) over two months, and users could reproduce the dam dataset by uploading the SWB polygons in the historical satellite images in 2019 in Google Earth Pro and following the procedure described here. Note that satellite images in the Google Earth Pro may change over time, but such change may have very limited impact on identification of dams as dams

254   have often stayed for many years after their construction.

2.4 Cross-comparison with other lake and reservoir datasets

256   To better understand the improvements and potential applications of our China-LDRL dataset,

257   we compared it with other three available dam and reservoir datasets: the GOODD, GRanD V1.3,

258   and GeoDAR datasets (**Fig. 2**). We first compared the dam quantity and areas of large reservoirs

259   at the provincial and national scales. Then, we checked the spatial distribution of each dam from

260   these datasets within Google Earth imagery as all these datasets provide detailed georeferenced

261   coordinates for some of dams, and the georeferenced information could be directly acquirable from

262   the spatial longitude and latitude. Here we did not compare the reservoir area with the GOODD

263   dataset as it does not provide such attribute except for catchment area (**Fig. 2d**).

3. Results

3.1 Annual map of surface water bodies in China for 2019

266   Surface water body frequency (F_{SWB}) of individual pixels for 2019 varied substantially across

267   China (**Fig. 4**). There were ~3.38 million seasonal surface water pixels (30-m spatial resolution)

268   in China, amounting to ~$3,375.88 \times 10^3$ km^2 seasonal surface water area (SWA) in 2019. Xinjiang

269   Province had the largest seasonal SWA (751.14×10^3 km^2), followed by Tibet (600.70×10^3 km^2),

270   Qinghai (564.57×10^3 km^2), Inner Mongolia (511.42×10^3 km^2), and Heilongjiang Province

271   (343.33×10^3 km^2) (**Fig. 5a**). There were ~0.21 million yearlong surface water pixels in China for

272   2019, amounting to ~214.92×10^3 km^2 yearlong SWA, which were mainly located in Tibet (62.65

273 $\times 10^3$ km^2), Qinghai (41.08 $\times 10^3$ km^2), and Xinjiang (24.60 $\times 10^3$ km^2) Provinces (**Fig. 5b**).

274 Additionally, Heilongjiang, Jiangsu, Inner Mongolia, Hubei, and Anhui Provinces also had

275 relatively larger yearlong SWA (> 5 $\times 10^3$ km^2) than other provinces in China.

276 **Insert Fig. 5 here**

277 **3.2 Numbers and areas of yearlong surface water bodies with different sizes in China**

278 The numbers and areas of yearlong SWB polygons of different sizes in China differed

279 considerably for 2019 (**Fig. 6**). In terms of yearlong SWB numbers, out of a total of 3.52 $\times 10^6$

280 yearlong SWB polygons in China in 2019, approximate 3.51 $\times 10^6$ polygons (99.57%) had an area

281 of ≤ 1 km^2, and ~2.16 $\times 10^6$ polygons (61.19%) had an area of ≤ 0.0036 km^2 (covering only 2 $\times$ 2

282 Landsat grid cells). Only 15 $\times 10^3$ (0.43%) yearlong SWB polygons had an area of > 1 km^2, and

283 359 polygons had an area of > 100 km^2. In terms of yearlong SWB areas, out of a total of 214.92

284 $\times 10^3$ km^2 yearlong SWA in China in 2019, large SWB polygons (size > 1 km^2) accounted for

285 83.54%, and very large SWB polygons (size > 100 km^2) accounted for 52.48%.

286 The numbers and areas of yearlong SWB polygons of different sizes at the provincial scale

287 had similar distribution patterns with those at the national scale (**Fig. S4, S5**). Almost all the

288 yearlong SWB polygons in individual provinces had an area of ≤ 1 km^2 (**Fig. S4**), however, those

289 SWB polygons with an area of > 1 km^2 accounted for a large proportion of SWA in most provinces

290 (**Fig. S5**). Those yearlong SWB polygons with an area of > 100 km^2 were mostly very large lakes

291 and rivers, which were mainly located in Tibet, Xinjiang, Qinghai, Jiangxi, and Heilongjiang

292   Provinces (**Fig. S5**) (Feng et al., 2019). Some provinces also had very large-size reservoirs, such

293   as Miyun Reservoir in Beijing, whose polygon size was greater than 100 km^2.

294   **Insert Fig. 6 here**

295   **3.3 Numbers, areas, and distribution of large dams, reservoirs, and lakes in China**

296   We identified 2,418 large dams in China, including 624 off-stream dams and 1,794 on-stream

297   dams, most of which were located in South, East, and Northeast China, as well as Xinjiang of

298   Northwest China (**Fig. 7a**). At the provincial scale, Xinjiang had the largest number of off-stream

299   dams (67), followed by Shandong (62), Heilongjiang (46), and Anhui (45) Provinces. Three

300   provinces (Hubei, Yunnan, and Guangdong) also had relatively larger off-stream dam numbers ($\geq$

301   40) than other provinces. Chongqing, Qinghai, and Tibet had no off-stream dams (**Fig. 7b**). Most

302   of on-stream dams in China were distributed in those provinces with large rivers. Guangdong had

303   the largest number of on-stream dams (172) in China, followed by Hubei (146), Heilongjiang (132),

304   Shandong (112), Jilin (103), and Sichuan (103) Provinces (**Fig. 7c**). However, there were no on-

305   stream dams in Shanghai. In terms of the functions of two kinds of dams and the spatial patterns

306   of climate (e.g. precipitation, temperature) and social-economic factors (e.g. population, GDP,

307   irrigation area) in South and North China, the provinces in Northeast and East China had larger

308   percentage of off-stream dams, whereas the provinces in Northeast and South China had larger

309   percentage of on-stream dams (**Fig. 7d**).

310   **Insert Fig. 7 here**

311   China had 3,051 large lakes with an area of > 1 km^2 in 2019, most of which were distributed

312   in West China, the Lower Yangtze River Basin, and Northeast China (**Fig. 8a, S6**), and they

313   together amounted to ~73.38 × 10^3 km^2. Tibet in West China had the largest number of lakes (966),

314   followed by Qinghai (479), Xinjiang (350), Inner Mongolia (234), and Heilongjiang (174)

315   Provinces (**Fig. 8b**). The lake areas in China had similar spatial patterns with the lake numbers

316   (**Fig. 8c**), and the western provinces in China had much larger lake areas than other provinces,

317   especially Tibet and Qinghai Provinces with 31.73 × 10^3 km^2 and 15.78 × 10^3 km^2, respectively.

318   As reservoirs and dams usually exist simultaneously, the spatial patterns of reservoir numbers and

319   areas matched well with those of dam numbers (**Figs. 7b, 8e-f**). In total, China had 2,194 large

320   reservoirs in 2019, they together amounted to an area of ~16.35 × 10^3 km^2. Xinjiang in Northwest

321   China had the largest reservoir area (1,923.11 km^2), followed by Heilongjiang (1,468.48 km^2),

322   Jiangsu (1,309.95 km^2), and Hubei (1,190.75 km^2) Provinces. In contrast, Tibet (18.34 km^2),

323   Shanghai (36.61 km^2), and Ningxia (45.40 km^2) had much smaller reservoir areas than other

324   provinces in China. In general, most of the dams and reservoirs in China were distributed in South

325   China, East China, and Northeast China, whereas most of lakes were located in West China, the

326   Lower Yangtze River Basin, and Northeast China.

327   **Insert Fig. 8 here**

4. Discussion

4.1 Improvements of the dataset of large dams, reservoirs, and lakes in China

In order to validate the reliability of our China-LDRL dataset, we first compared the numbers of large dams and areas of large reservoirs between our dataset and published datasets (GOODD, GRanD, and GeoDAR), then we checked the geographical coordinates of dams within the historical satellite images in 2019 in Google Earth Pro.

The GOODD dataset has the largest number of dams (9,234) in China among these published global datasets (**Fig. 2a**). However, it includes both large, moderate, and small dams, and does not report the corresponding reservoir attributes (e.g. reservoir area), which limits its applications to water-related research (Paredes-Beltran et al., 2021). The GRanD dataset has the smallest number (814) of large dams with reservoir area > 1 km^2 in China (**Fig. 9b, e**) as the dam information was provided by multiple institutions from the world (Lehner et al., 2011), which clearly underestimates the number of dams. The GeoDAR dataset has a larger number of large dams (1,162) than the GRanD dataset, because it was generated by combining the GRanD and ICOLD WRD datasets (Wang et al., 2021a). However, our China-LDRL dataset identified 2,418 large dams and 2,194 large reservoirs (**Fig. 9d, e**), making substantial improvement of large dam and reservoir dataset in China.

The number differences of large dams between our China-LDRL and the GRanD and GeoDAR datasets could be explained by several factors. First, our study used all the available

347   Landsat images in 2019 and a more accurate SWB mapping algorithm to generate SWB maps in

348   China, however, the GRanD and GeoDAR datasets used the SWBD map (produced in 2000)

349   (Slater et al., 2006) and the surface water maps during 1984-2015 produced by the JRC (Pekel et

350   al., 2016). We were able, therefore, to integrate more Landsat images and get more SWB polygons,

351   as well as larger numbers of large dams and reservoirs than other datasets. In addition, the different

352   strategies for identifying dams also caused the differences of dam numbers. The dam information

353   from the GRanD dataset was contributed by eleven participating institutions, and the GeoDAR

354   dataset combined two published dam datasets (WRD and GRanD), rechecked detailed dam

355   information using the Google Maps geocoding API, and then reported the georeferenced

356   information of dams. Unlike the GRanD and GeoDAR datasets, our study first generated SWB

357   raster and vector maps using the mNDWI/VIs SWB mapping algorithm, and then selected the large

358   yearlong SWB polygons with area > 1 km^2. After that, we visually checked the large SWB

359   polygons one by one and identified each dam with accurate geographical coordinates (**Fig. S3**). In

360   addition to the dam numbers, we also compared the reservoir areas between different datasets (**Fig.**

361   **S7**). Our China-LDRL dataset reported ~16.35 $\times$ 10^3 km^2 large reservoir area, which was smaller

362   than those of the GRanD (20.98 $\times$ 10^3 km^2) and GeoDAR (21.84 $\times$ 10^3 km^2) datasets. The GRanD

363   v1.3 dataset linked the "maximum surface water extent" from the JRC dataset to the corresponding

364   dams as the reservoir regions, however, we used the "yearlong surface water body" to depict the

365   reservoirs in the China-LDRL dataset, which might have made our reservoir areas smaller (**Fig.**

366   **S8**).

367 **Insert Fig. 9 here**

368   In this study, we also checked the accuracy of geographical coordinates of dams from these

369   dam datasets. Here we first uploaded above-mentioned three dam datasets and our China-LDRL

370   dataset in the Google Earth Pro and visually checked the spatial distribution of each dam within

371   the historical satellite images in 2019 (**Fig. 10**). We found that the dam locations of the GOODD

372   dataset had substantial geographic offsets, some of which are larger than 500 m (**Fig. S9**). We

373   further overlapped the GOODD dam layer with our yearlong SWB map (Section 2.3.1), and the

374   results showed that only $12.52 \pm 3.87\%$ of the GOODD dams were intersected with the SWB layer

375   at the national scale (**Fig. S10a**). In the case that we applied a 100-m and 500-m tolerance when

376   intersecting the GOODD dams with our yearlong SWB map for 2019, the intersection rate

377   increased to only $47.58 \pm 9.70\%$ and $76.46 \pm 7.11\%$, respectively (**Figs. S10b, S11**). In addition,

378   we applied different tolerances when the GRanD and GeoDAR datasets intersected with our

379   yearlong SWB layer. About $65.57 \pm 6.79\%$ of the dams in the GRanD dataset intersected with our

380   yearlong SWB map, which increased to $87.52 \pm 6.45\%$ and $95.94 \pm 4.49\%$ when using a 100-m

381   and 500-m tolerance. Although the GeoDAR dataset is just updated and the newest version had

382   much higher accuracy than the previous version, its geographical coordinates also had some offsets

383   (**Fig. 10f, g**), and $58.49 \pm 6.07\%$ of its dams intersected with the yearlong SWB layer, and 82.33

384    $\pm 3.98\%$ and $90.22 \pm 3.18\%$ intersected when the tolerance was 100-m and 500-m. Different

385   methods and purposes caused the georeferenced offsets of these datasets. For example, the original

386   digitized dam points in GOODD V1.0 were purposefully snapped to the 30-arc-second

387   HydroSHEDS river networks, leading to the offset from the actual dam locations. On the other

388   hand, GOODD v1.0 is directly compatible with HydroSHEDS and is therefore more convenient

389   for modeling purposes. In GeoDAR v1.1, dam points in China were georeferenced using the

390   Google Maps geocoding API, and many dam labels fell on the reservoir surface instead of the

391   dams. Additionally, Google Maps in China have substantial misalignment (500 m to 1 km or so)

392   between the satellite images and the map labels due to China's GPS shift problem, resulting in the

393   geographic offsets even though the geocoding procedure is correct. In total, these comparisons

394   suggested the improved accuracy of our China-LDRL dataset, which could provide important and

395   reliable information for water resource management and water security in China.

396   **Insert Fig. 10 here**

397   **4.2 Uncertainties, limitations, outlooks, and implications**

398   In this study, we produced detailed and more accurate China's open surface water bodies, large

399   dams, reservoirs, and lakes (China-LDRL) dataset for 2019, and analyzed their spatial distribution

400   patterns. This study benefited from the usage of time-series Landsat imagery and GEE cloud

401   computing platform, as well as simple and robust SWB mapping algorithms. First, time series

402   Landsat images at high spatial resolution (30-m) provide larger numbers of good-quality

403   observations for identifying SWB. Second, GEE cloud computing platform enables us to acquire

404   and analyze tens of thousands of Landsat images in hours. Third, the mNDWI/VIs algorithm used

405   in this study reduced the uncertainties induced by the bad-quality observations and provide

406   accurate SWB maps. Finally, we visually checked the large SWB polygons (area > 1 km^2) one by

407 one by using the historical satellite images in 2019 within China in Google Earth Pro, and we

408 recorded the georeferenced coordinates of individual dams in China for 2019.

409   We would also acknowledge that the data quality of input satellite images remains to be a

410 concern for the identification of dams, reservoirs, and lakes. The spatial distribution of good-

411 quality observations of Landsat data shows that more than 98.36% of the total 30-m pixels in China

412 had more than 5 good-quality observations and more than 91.24% of the total pixels had more than

413 10 good-quality observations for 2019 (**Fig. 1b**), but the regions with complex topography and

414 mountains, such as South and Southwest China, had much fewer good-quality observations than

415 other regions, which might underestimate surface water areas, as well as dam and reservoir

416 numbers and areas. In addition, it is impossible to remove all the bad-quality observations (e.g.

417 clouds, terrain shadows) because of the limited quality of the QA band and digital elevation model

418 data in GEE. Therefore, the remaining bad-quality observations could result in some inevitable

419 uncertainties in the resultant maps. In the future, as more images from Landsat dataset and other

420 high spatial resolution sensors (e.g., Sentinel-1, Sentinel-2) are added into GEE platform (Wulder

421 et al., 2016), SWB mapping accuracy could be further improved, providing more detailed

422 geospatial data of dams, reservoirs, and lakes in China. In addition, visual interpretation method

423 for identifying dams and reservoirs in this study could also bring about some uncertainties to the

424 classification of dams/reservoirs due to the limitations of knowledge and experience of interpreters,

425 such as the misclassification of some reservoirs regulated by dams/gates as lakes (e.g. Hongze

426 Lake in Jiangsu Province) and the misclassification between on-stream and off-stream

427 dams/reservoirs.

428   In our China-LDRL dataset, we identified and reported those large SWB, however, the

429 importance of monitoring small water bodies (area $\leq$ 1 km^2) and dams is gradually recognized as

430 they play critical roles in accurate assessments of their agricultural potential or their cumulative

431 influence on watershed hydrology (Ogilvie et al., 2018). In the future, we can include these small

432 SWB polygons into our dataset to enhance the spatial details and distributions of dams, reservoirs,

433 and lakes in China.

434   The conversions between rivers, lakes, and reservoirs have critical effects on the ecosystem

435 services. For example, the construction of the Three Gorges Dam contributed to the decrease of

436 surface water area and biodiversity in its downstream areas (Fang et al., 2006; Feng et al., 2013;

437 Wang et al., 2020a), and reduced the sediment loads in the Yangtze River, causing the decreased

438 deposition rates of coastal wetlands in the Yangtze Delta (Feng et al., 2016; Wang et al., 2021b).

439 Furthermore, the conversion from natural lakes and rivers to man-made reservoirs has

440 disproportionate effects on the local, regional, and global carbon cycle (Howard Coker et al., 2009).

441 For example, dam construction has reduced the areal extent of CO_2 gas exchange in natural rivers

442 (Ran et al., 2021). In the future, more detailed information (e.g. construction year of dam) needs

443 to be included in our China-LDRL dataset, making it possible to analyze the effects of conversions

444 from natural lakes and rivers to reservoirs on the biodiversity and carbon cycle.

445

5. Data availability

The China-LDRL dataset is publicly available at https://doi.org/10.6084/m9.figshare.16964656.v3 (Wang et al., 2022), and it includes three shapefiles. The "China_large_dams.shp" is the large dams in China with five attributes, including ID (dam_id), dam class (dam_class, "1" means on-stream dam and "-1" means off-stream dam), longitude, latitude, and corresponding reservoir ID (reser_id). The "China_large_lakes.shp" is the large lakes map in China with three attributes: ID, lake area (poly_area, km^2), and lake perimeter (poly_len, km). The "China_large_reservoirs.shp" is the large reservoirs map in China with five attributes: ID, reservoir area (poly_area, km^2), reservoir perimeter (poly_len, km), corresponding dam ID (dam_ID), and dam classes (dam_class).

6. Code availability

Code used in calculations of surface water bodies is available upon request.

7. Conclusion

Several studies have published global or national dam, reservoir, and lake datasets based on satellite images (**Table 1**). However, these datasets usually have large georeferenced coordinate offsets, which poses some limitations to those studies that aim to address major issues in hydrology, ecology, and water resource management in China. In this study, we have generated the dataset of China's open surface water bodies, large dams, reservoirs, and lakes (China-LDRL) for 2019, and

then analyzed their spatial distributions at the provincial and national scales. Satellite image data quality is still a major source of uncertainty that affects the accuracy of the surface water body maps. As more images from Landsat datasets and other high spatial resolution sensors (e.g., Sentinel-1, Sentinel-2) are added to GEE platform, the accuracy of SWB maps can be further improved, providing more detailed geospatial data of dams, reservoir, and lakes in China. The provision of the reliable, accurate China-LDRL dataset on dams, reservoirs, and lakes will contribute to the understanding of water resources management and water security in China.

Author contributions

X.X., X.W., and B.L. designed the study. X.W. carried out image data processing and led interpretation of the results and writing of the manuscript. Y.Q., and J.D. contributed to image data processing, X.X., B.L., Y.Q., J.D., and J. W. contributed to the interpretation and discussion of the results.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

This study was supported in part by research grants from the U.S. National Science Foundation

483 (1911955), the Natural Science Foundation of China (81961128002), the China Postdoctoral

484 Science Foundation (2021M700835 and 2021TQ0072), and the China Scholarship Council

485 (201906100124).

References

487 Armstrong, A.: Lake carbon, Nat. Geosci., 3, 151, https://doi.org/10.1038/ngeo816, 2010.

488 Birkett, C. M. and Mason, I. M.: A New Global Lakes Database for a Remote Sensing Program

489 Studying Climatically Sensitive Large Lakes, J. Great Lakes Res., 21,

490 https://doi.org/10.1016/S0380-1330(95)71041-3, 1995.

491 Dalin, C., Hanasaki, N., Qiu, H., Mauzerall, D. L., and Rodriguez-Iturbe, I.: Water resources

492 transfers through Chinese interprovincial and foreign food trade, Proc. Natl. Acad. Sci. U. S. A.,

493 111, 9774–9779, https://doi.org/10.1073/pnas.1404749111, 2014.

494 Fang, J., Wang, Z., Zhao, S., Li, Y., Tang, Z., Yu, D., Ni, L., Liu, H., Xie, P., Da, L., Li, Z., and

495 Zheng, C.: Biodiversity changes in the lakes of the Central Yangtze,

496 https://doi.org/10.1890/1540-9295(2006)004[0369:BCITLO]2.0.CO;2, 2006.

497 Feng, L., Hu, C., Chen, X., and Zhao, X.: Dramatic inundation changes of China's two largest

498 freshwater lakes linked to the Three Gorges Dam, Environ. Sci. Technol., 47, 9628–9634,

499 https://doi.org/10.1021/es4009618, 2013.

500 Feng, L., Han, X., Hu, C., and Chen, X.: Four decades of wetland changes of the largest

501   freshwater lake in China: Possible linkage to the Three Gorges Dam?, Remote Sens. Environ.,

502   176, 43–55, https://doi.org/10.1016/j.rse.2016.01.011, 2016.

503   Feng, S., Liu, S., Huang, Z., Jing, L., Zhao, M., Peng, X., Yan, W., Wu, Y., Lv, Y., Smith, A. R.,

504   McDonald, M. A., Patil, S. D., Sarkissian, A. J., Shi, Z., Xia, J., and Ogbodo, U. S.: Inland water

505   bodies in China: Features discovered in the long-term satellite data, Proc. Natl. Acad. Sci. U. S.

506   A., 116, 25491–25496, https://doi.org/10.1073/pnas.1910872116, 2019.

507   Gao, H.: Satellite remote sensing of large lakes and reservoirs: from elevation and area to

508   storage, Wiley Interdiscip. Rev. Water, 2, 147–157, https://doi.org/10.1002/wat2.1065, 2015.

509   Gao, H., Birkett, C., and Lettenmaier, D. P.: Global monitoring of large reservoir storage from

510   satellite remote sensing, Water Resour. Res., 48, W09504,

511   https://doi.org/10.1029/2012WR012063, 2012.

512   Hayes, N. M., Deemer, B. R., Corman, J. R., Razavi, N. R., and Strock, K. E.: Key differences

513   between lakes and reservoirs modify climate signals: A case for a new conceptual model,

514   Limnol. Oceanogr. Lett., 2, 47–62, https://doi.org/https://doi.org/10.1002/lol2.10036, 2017.

515   Howard Coker, E., Hotchkiss, R. H., and Johnson, D. A.: Conversion of a Missouri River Dam

516   and Reservoir to a Sustainable System: Sediment Management1, JAWRA J. Am. Water Resour.

517   Assoc., 45, 815–827, https://doi.org/https://doi.org/10.1111/j.1752-1688.2009.00324.x, 2009.

518   ICOLD: World register of dams, Paris, 2011.

519   Lehner, B. and Döll, P.: Development and validation of a global database of lakes, reservoirs and

520   wetlands, J. Hydrol., 296, 1–22, https://doi.org/10.1016/j.jhydrol.2004.03.028, 2004.

521   Lehner, B., Liermann, C. R., Revenga, C., Vörömsmarty, C., Fekete, B., Crouzet, P., Döll, P.,

522   Endejan, M., Frenken, K., Magome, J., Nilsson, C., Robertson, J. C., Rödel, R., Sindorf, N., and

523   Wisser, D.: High-resolution mapping of the world's reservoirs and dams for sustainable river-

524   flow management, Front. Ecol. Environ., 9, 494–502, https://doi.org/10.1890/100125, 2011.

525   Liu, J. and Yang, W.: Water sustainability for China and beyond, Science (80-.)., 337, 649–650,

526   https://doi.org/10.1126/science.1219471, 2012.

527   Ma, R. H., Yang, G. S., Duan, H. T., Jiang, J. H., Wang, S. M., Feng, X. Z., Li, A. N., Kong, F.

528   X., Xue, B., Wu, J. L., and Li, S. J.: China's lakes at present: Number, area and spatial

529   distribution, Sci. China Earth Sci., 54, 283–289, https://doi.org/10.1007/s11430-010-4052-6,

530   2011.

531   Mulligan, M., van Soesbergen, A., and Sáenz, L.: GOODD, a global dataset of more than 38,000

532   georeferenced dams, Sci. Data, 7, 1–8, https://doi.org/10.1038/s41597-020-0362-5, 2020.

533   Murray, N. J., Phinn, S. R., DeWitt, M., Ferrari, R., Johnston, R., Lyons, M. B., Clinton, N.,

534   Thau, D., and Fuller, R. A.: The global distribution and trajectory of tidal flats, Nature, 565, 222,

535   https://doi.org/10.1038/s41586-018-0805-8, 2019.

536   Ogilvie, A., Belaud, G., Massuel, S., Mulligan, M., Le Goulven, P., and Calvez, R.: Surface

537   water monitoring in small water bodies: Potential and limits of multi-sensor Landsat time series,

Hydrol. Earth Syst. Sci., 22, 4349–4380, https://doi.org/10.5194/hess-22-4349-2018, 2018.

Paredes-Beltran, B., Sordo-Ward, A., and Garrote, L.: Dataset of Georeferenced Dams in South America (DDSA), Earth Syst. Sci. Data, https://doi.org/10.5194/essd-13-213-2021, 2021.

Pekel, J. F., Cottam, A., Gorelick, N., and Belward, A. S.: High-resolution mapping of global surface water and its long-term changes, Nature, 540, 418, https://doi.org/10.1038/nature20584, 2016.

Ran, L., Butman, D. E., Battin, T. J., Yang, X., Tian, M., Duvert, C., Hartmann, J., Geeraert, N., and Liu, S.: Substantial decrease in CO_2 emissions from Chinese inland waters due to global change, Nat. Commun., 12, 1–9, https://doi.org/10.1038/s41467-021-21926-6, 2021.

Slater, J. A., Garvey, G., Johnston, C., Haase, J., Heady, B., Kroenung, G., and Little, J. K.: The SRTM Data Finishing Process and Products, Photogramm. Eng. Remote Sensing, 72, 237–247, 2006.

Tao, S., Zhang, H., Feng, Y., and Zhu, J.: Changes in China's water resources in the early 21st century, Front. Ecol. Environ., https://doi.org/10.1002/fee.2164, 2020.

Thornton, J., Steel, A., and Rast, W.: Chapter 8* - Reservoirs, in: Water quality assessments : a guide to the use of biota, sediments and water in environmental monitoring, 2–3, 1996.

Wang, J., Sheng, Y., Gleason, C. J., and Wada, Y.: Downstream Yangtze River levels impacted by Three Gorges Dam, Environ. Res. Lett., 8, 044012, https://doi.org/10.1088/1748-

556   9326/8/4/044012, 2013.

557   Wang, J., Walter, B. A., Yao, F., Song, C., Ding, M., Maroof, A. S., Zhu, J., Fan, C., Xin, A.,

558   McAlister, J. M., Sikder, S., Sheng, Y., Allen, G. H., Crétaux, J.-F., and Wada, Y.: GeoDAR:

559   Georeferenced global dam and reservoir dataset for bridging attributes and geolocations, Earth

560   Syst. Sci. Data, 2021, 1–52, https://doi.org/10.5194/essd-2021-58, 2021a.

561   Wang, X., Xiao, X., Zou, Z., Dong, J., Qin, Y., Doughty, R. B., Menarguez, M. A., Chen, B.,

562   Wang, J., Ye, H., Ma, J., Zhong, Q., Zhao, B., and Li, B.: Gainers and losers of surface and

563   terrestrial water resources in China during 1989 – 2016, Nat. Commun., 11, 1–12,

564   https://doi.org/10.1038/s41467-020-17103-w, 2020a.

565   Wang, X., Xiao, X., Zou, Z., Hou, L., Qin, Y., Dong, J., Doughty, R. B., Chen, B., Zhang, X.,

566   Chen, Y., Ma, J., Zhao, B., and Li, B.: Mapping coastal wetlands of China using time series

567   Landsat images in 2018 and Google Earth Engine, ISPRS J. Photogramm. Remote Sens., 163,

568   312–326, https://doi.org/10.1016/j.isprsjprs.2020.03.014, 2020b.

569   Wang, X., Xiao, X., Xu, X., Zou, Z., Chen, B., Qin, Y., Zhang, X., Dong, J., Liu, D., Pan, L., and

570   Li, B.: Rebound in China's coastal wetlands following conservation and restoration, Nat.

571   Sustain., 4, 1076–1083, https://doi.org/10.1038/s41893-021-00793-5, 2021b.

572   Wu, J., Huang, J., Han, X., Gao, X., He, F., Jiang, M., Jiang, Z., Primack, R. B., and Shen, Z.:

573   The Three Gorges Dam: An ecological perspective, Front. Ecol. Environ., 2, 241–248,

574   https://doi.org/10.1890/1540-9295(2004)002[0241:TTGDAE]2.0.CO;2, 2004.

575 Wulder, M. A., White, J. C., Loveland, T. R., Woodcock, C. E., Belward, A. S., Cohen, W. B.,

576 Fosnight, E. A., Shaw, J., Masek, J. G., and Roy, D. P.: The global Landsat archive: Status,

577 consolidation, and direction, Remote Sens. Environ., 185, 271–283,

578 https://doi.org/10.1016/j.rse.2015.11.032, 2016.

579 Xiang, Y., Fu, Z., Meng, Y., Zhang, K., and Cheng, Z.: Analysis of wave clipping effects of

580 plain reservoir artificial islands based on MIKE21 SW model, Water Sci. Eng., 12, 179–187,

581 https://doi.org/https://doi.org/10.1016/j.wse.2019.08.002, 2019.

582 Yang, X. and Lu, X.: Drastic change in China's lakes and reservoirs over the past decades, Sci.

583 Rep., 4, 6041, https://doi.org/10.1038/srep06041, 2014.

584 Yang, X. and Lu, X. X.: Delineation of lakes and reservoirs in large river basins: An example of

585 the Yangtze River Basin, China, https://doi.org/10.1016/j.geomorph.2013.02.018, 2013.

586 Zhang, G., Yao, T., Chen, W., Zheng, G., Shum, C. K., Yang, K., Piao, S., Sheng, Y., Yi, S., Li,

587 J., Oreilly, C., Qi, S., Shen, S., Zhang, H., and Jia, Y.: Regional differences of lake evolution

588 across China during 1960s-2015 and its natural and anthropogenic causes, Remote Sens.

589 Environ., 221, 386–404, https://doi.org/10.1016/j.rse.2018.11.038, 2019.

590 Zhang, Q., Li, L., Wang, Y. G., Werner, A. D., Xin, P., Jiang, T., and Barry, D. A.: Has the

591 Three-Gorges Dam made the Poyang Lake wetlands wetter and drier?, Geophys. Res. Lett., 39,

592 L20402, https://doi.org/10.1029/2012GL053431, 2012.

593 Zhou, F., Bo, Y., Ciais, P., Dumas, P., Tang, Q., Wang, X., Liu, J., Zheng, C., Polcher, J., Yin,

594 Z., Guimberteau, M., Peng, S., Ottle, C., Zhao, X., Zhao, J., Tan, Q., Chen, L., Shen, H., Yang,

595 H., Piao, S., Wang, H., and Wada, Y.: Deceleration of China's human water use and its key

596 drivers, Proc. Natl. Acad. Sci. U. S. A., 117, 7702–7711,

597 https://doi.org/10.1073/pnas.1909902117, 2020.

598 Zhou, Y., Dong, J., Xiao, X., Xiao, T., Yang, Z., Zhao, G., Zou, Z., and Qin, Y.: Open surface

599 water mapping algorithms: A comparison of water-related spectral indices and sensors, 9, 256,

600 https://doi.org/10.3390/w9040256, 2017.

601 Zhou, Y., Dong, J., Liu, J., Metternicht, G., Shen, W., You, N., Zhao, G., and Xiao, X.: Are there

602 sufficient Landsat observations for retrospective and continuous monitoring of land cover

603 changes in China?, Remote Sens., 11, https://doi.org/10.3390/rs11151808, 2019a.

604 Zhou, Y., Dong, J., Xiao, X., Liu, R., Zou, Z., Zhao, G., and Ge, Q.: Continuous monitoring of

605 lake dynamics on the Mongolian Plateau using all available Landsat imagery and Google Earth

606 Engine, Sci. Total Environ., 689, 366–380, https://doi.org/10.1016/J.SCITOTENV.2019.06.341,

607 2019b.

608 Zhu, J., Song, C., Wang, J., and Ke, L.: China's inland water dynamics: The significance of

609 water body types, Proc. Natl. Acad. Sci., 117, 13876–13878,

610 https://doi.org/10.1073/pnas.2005584117, 2020.

611 Zhu, Z., Wang, S., and Woodcock, C. E.: Improvement and expansion of the Fmask algorithm:

612 Cloud, cloud shadow, and snow detection for Landsats 4-7, 8, and Sentinel 2 images, Remote

613   Sens. Environ., 159, 269–277, https://doi.org/10.1016/j.rse.2014.12.014, 2015.

614   Zou, Z. H., Dong, J. W., Menarguez, M. A., Xiao, X. M., Qin, Y. W., Doughty, R. B., Hooker,

615   K. V, and Hambright, K. D.: Continued decrease of open surface water body area in Oklahoma

616   during 1984-2015, Sci. Total Environ., 595, 451–460,

617   https://doi.org/10.1016/j.scitotenv.2017.03.259, 2017.

618   Zou, Z. H., Xiao, X. M., Dong, J. W., Qin, Y. W., Doughty, R. B., Menarguez, M. A., Zhang, G.

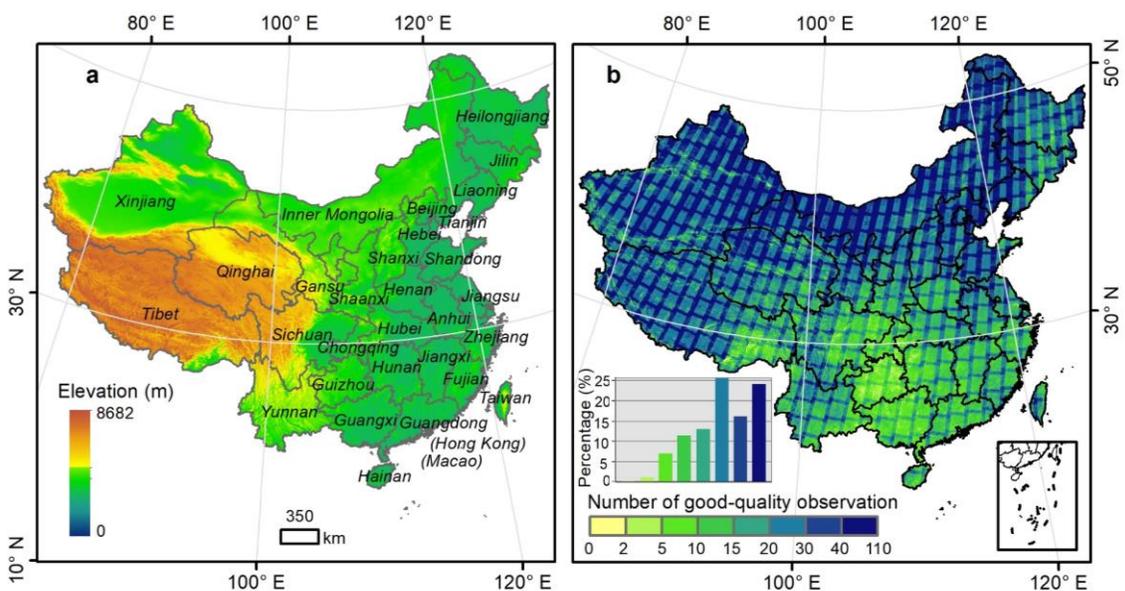
619   L., and Wang, J.: Divergent trends of open-surface water body area in the contiguous United

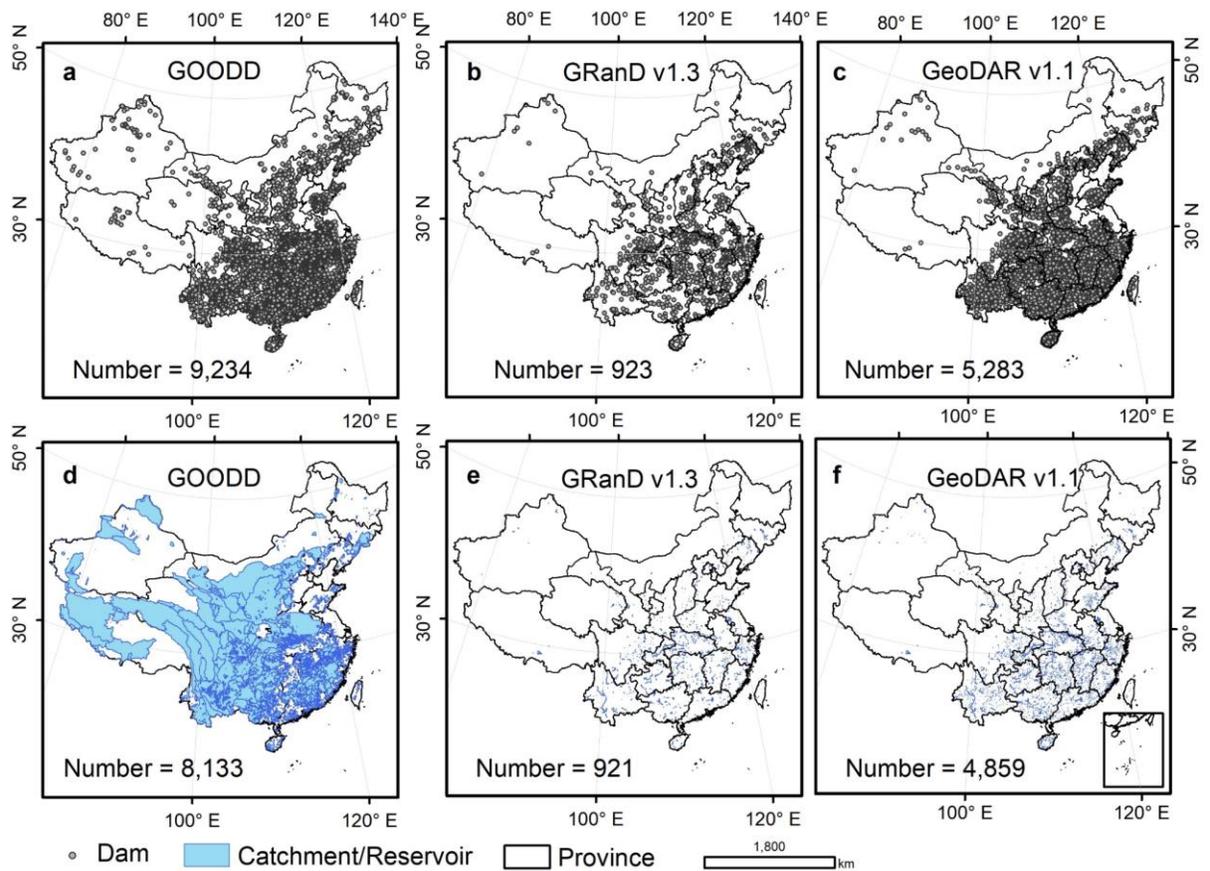
620   States from 1984 to 2016, Proc. Natl. Acad. Sci. U. S. A., 115, 3810–3815,

621   https://doi.org/10.1073/pnas.1719275115, 2018.

622

623

624   **Figures and Legends**

626   **Fig. 1.** Spatial distribution of provinces and elevation (a) and numbers of Landsat good-quality

627   observations (b) in China for 2019.

628

629 **Fig. 2**. Spatial distribution of dams from the GlObal GeOreferenced Database of Dams (GOODD)

630 (a), the Global Reservoir and Dam (GRanD) v1.3 (b), and the Georeferenced global Dam And

631 Reservoir (GeoDAR) v1.1 (c) datasets. The GOODD dataset reported the catchment of each dam

632 (d) while the GRanD and GeoDAR datasets reported the reservoir information of each dam (e, f).

633

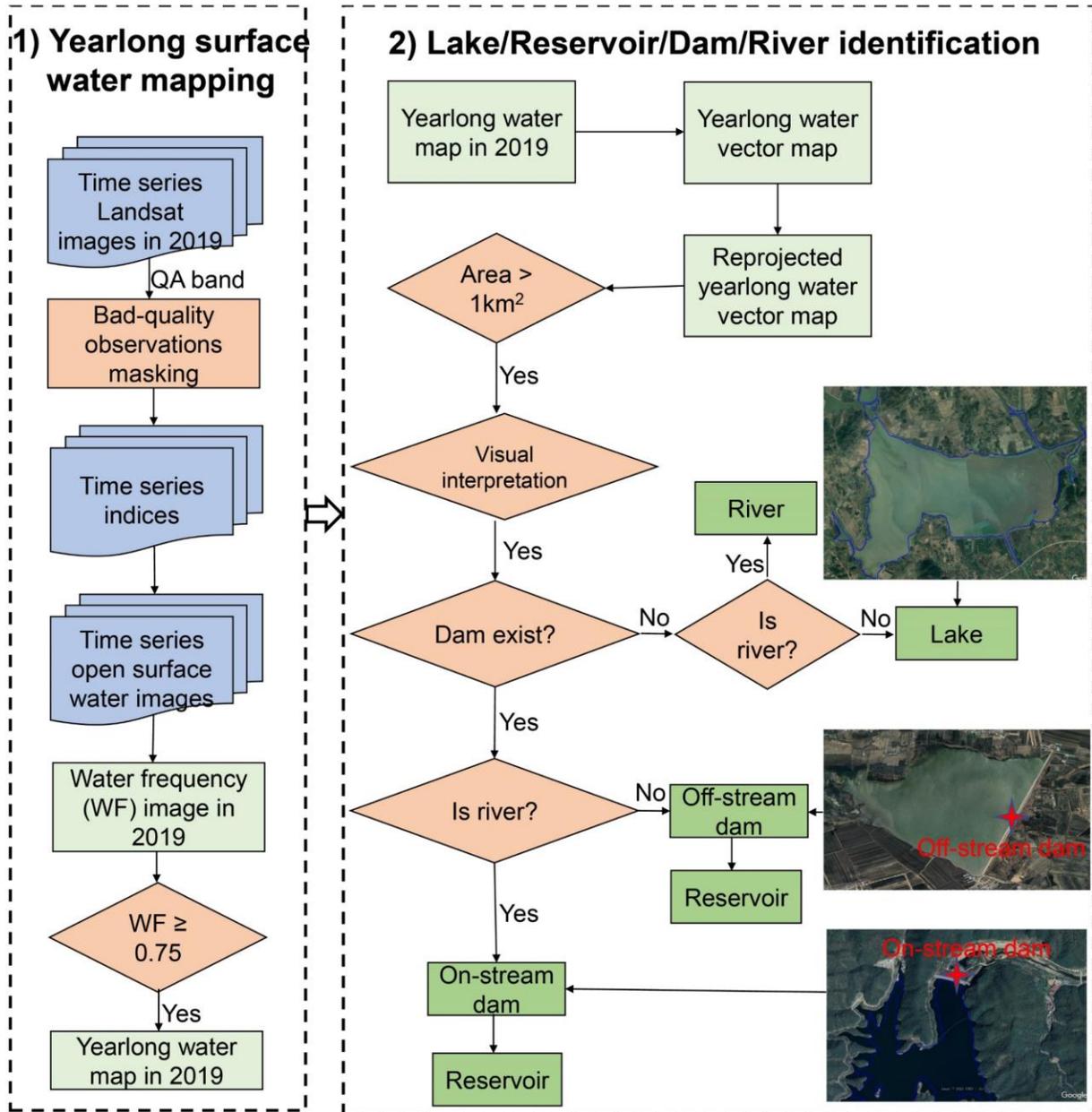



Fig. 3. Schematic flowchart of lakes, reservoirs, dams, and rivers identification in this study. The

images were acquired from Google Earth Pro (© Google Earth Pro 2019).

636

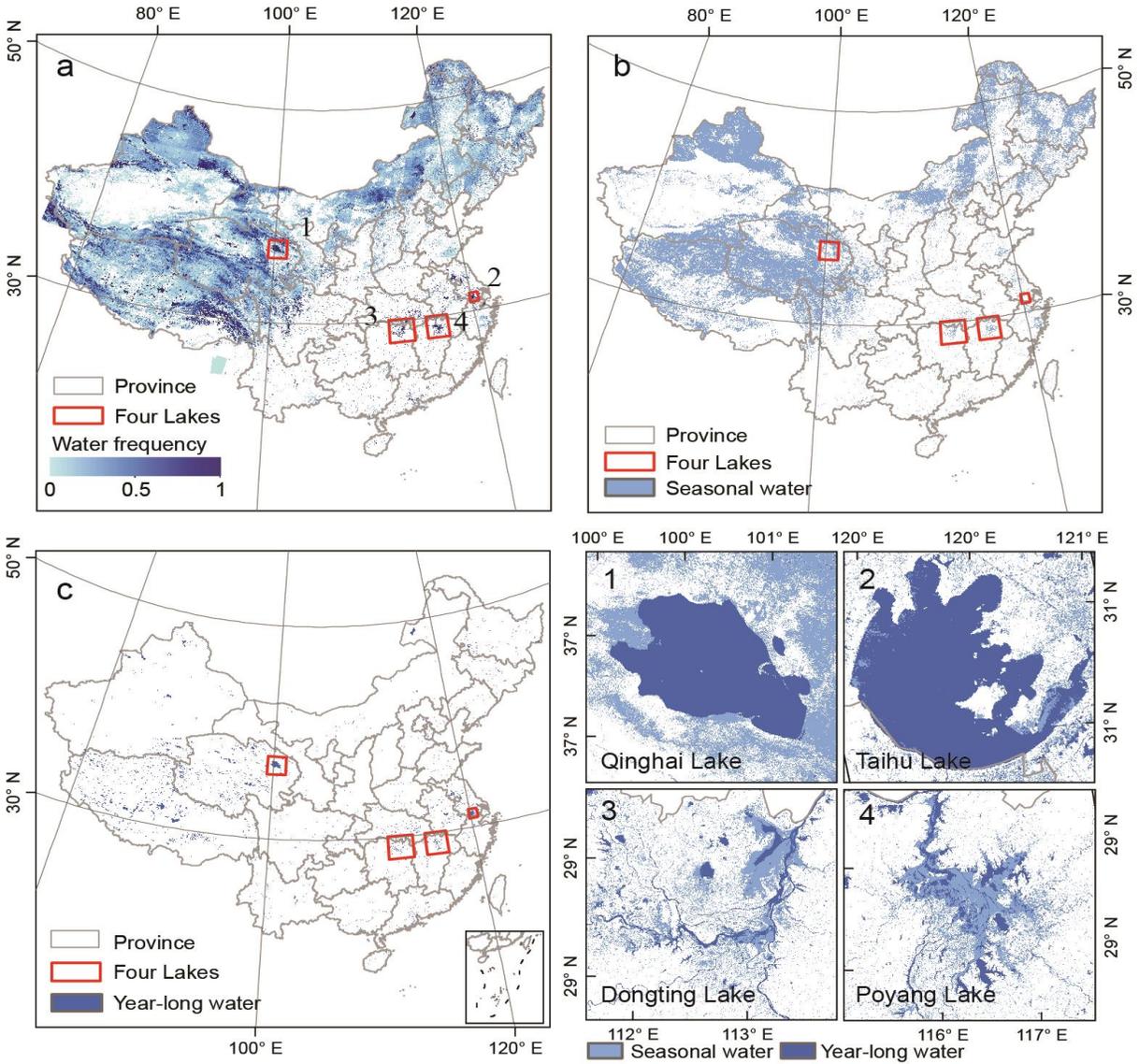
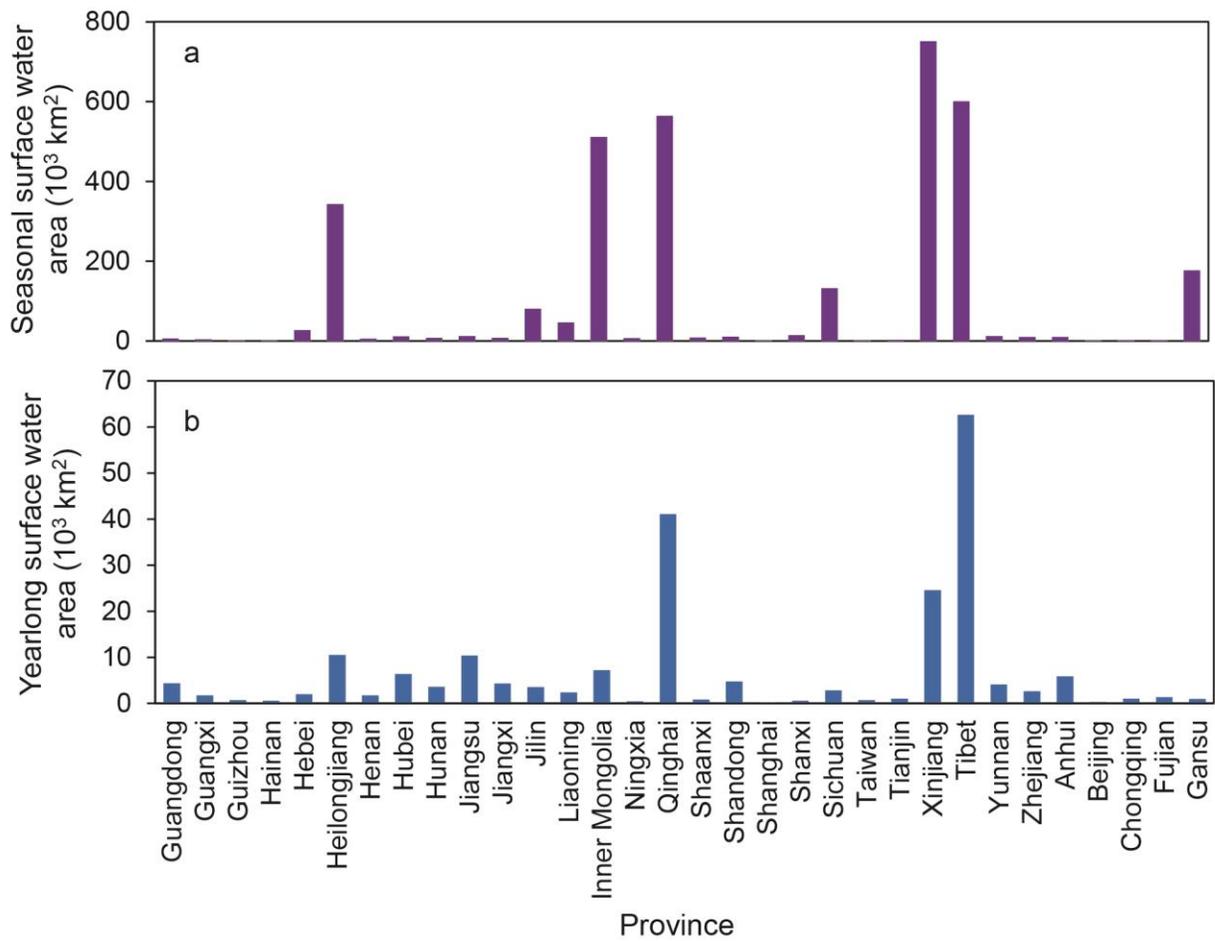
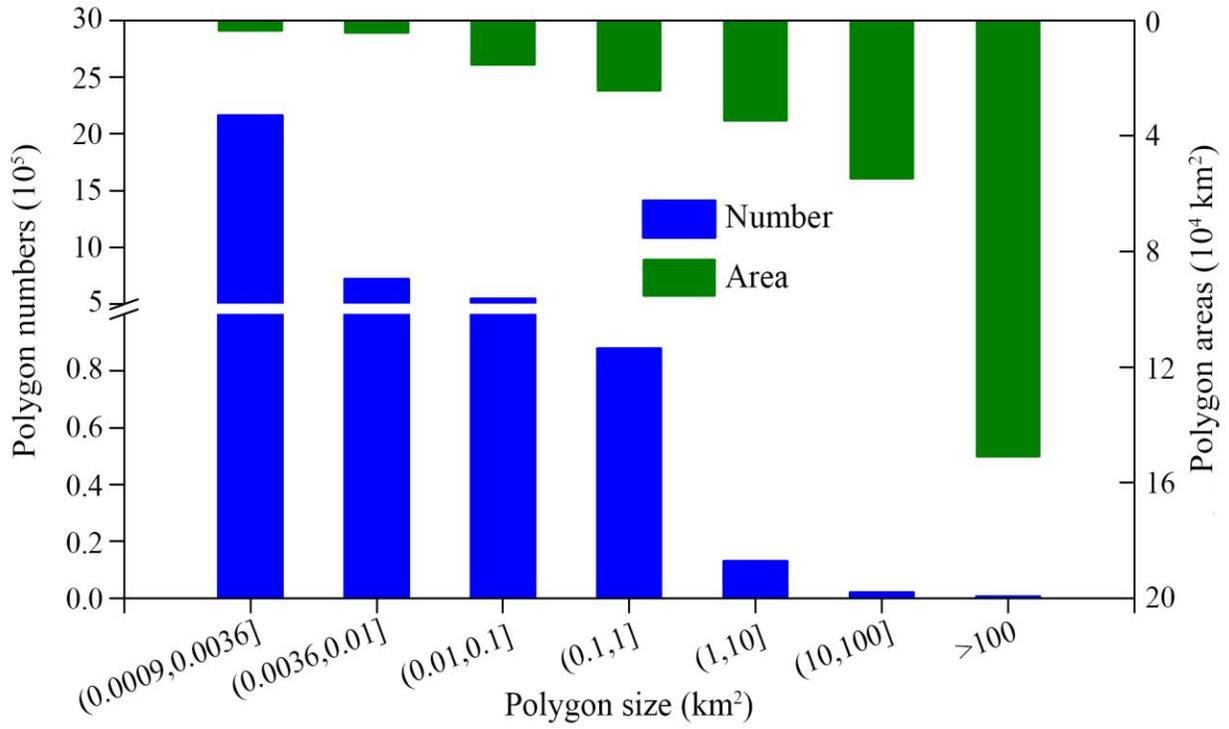



Fig. 4. Spatial distribution of surface water body (SWB) in China for 2019. (a), Water frequency,

(b), Seasonal SWB, (c), Yearlong SWB. Subfigures (1-4) are three zoom-in views of seasonal and

year-long SWB in Qinghai Lake, Taihu Lake, Dongting Lake, and Poyang Lake in China.

640

641   **Fig. 5.** Areas of seasonal (a) and yearlong (b) surface water bodies by province in China for 2019.

642

643 **Fig. 6.** Numbers and areas of yearlong surface water body polygons with different sizes.

644

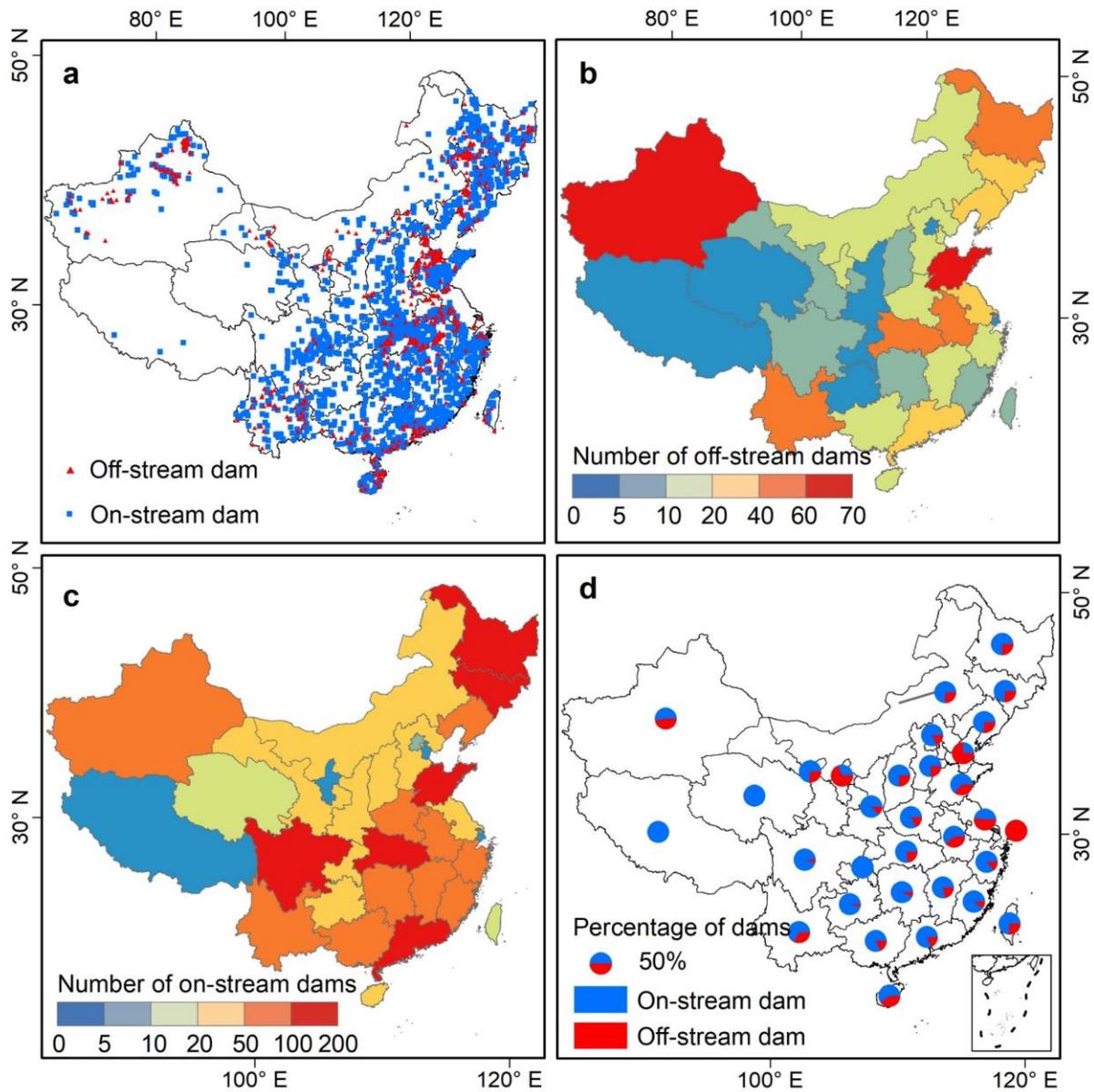



Fig. 7. Distribution of off-stream and on-stream dams in China for 2019. a, Spatial distribution of

dams; b, Number of off-stream dams by province; c, Number of on-stream dams by province; d,

Percentage of off-stream and on-stream dams by province.

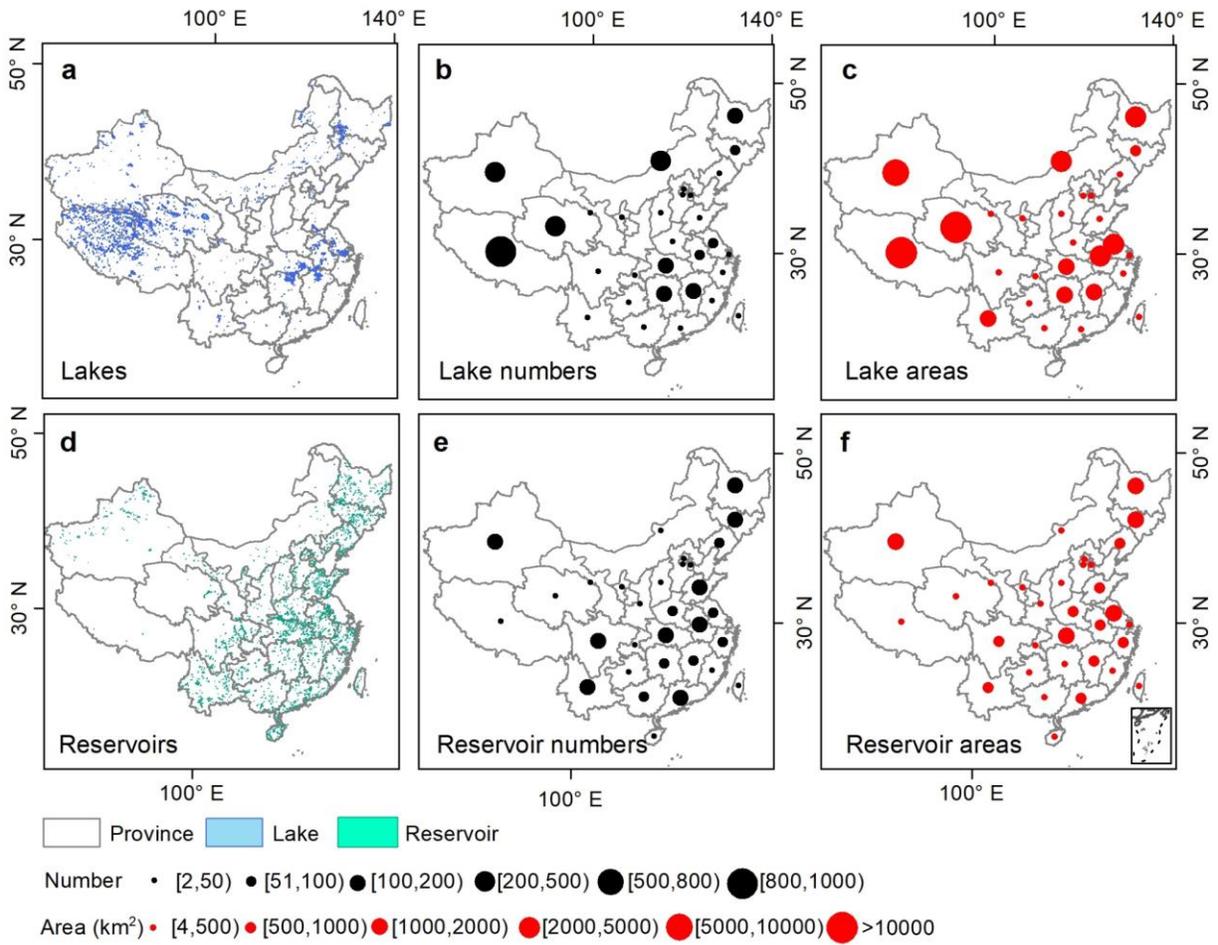



Fig. 8. Numbers and areas of lakes and reservoirs in China for 2019. a, Spatial distribution of lakes; b, Lake numbers by province; c, Lake areas by province; d, Spatial distribution of reservoirs; e, Reservoir numbers by province; f, Reservoir areas by province.

652

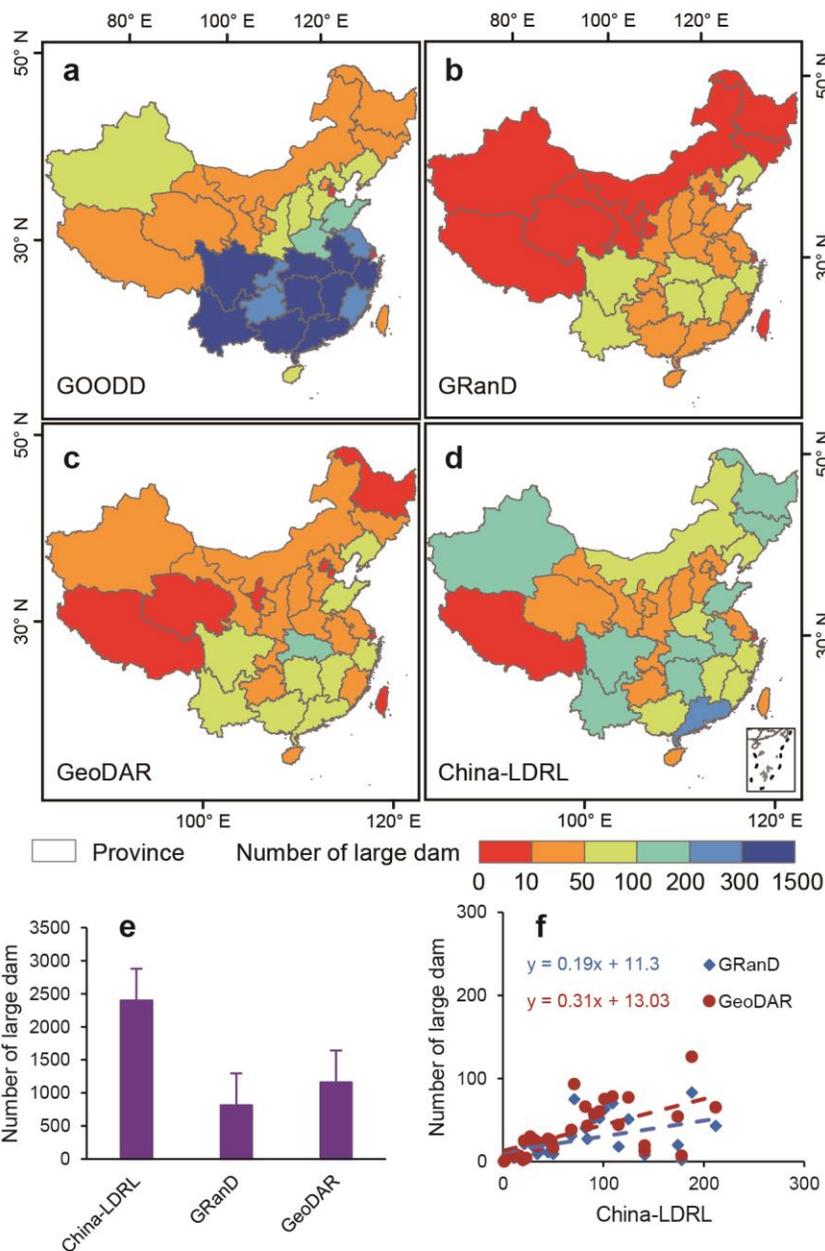



Fig. 9. Numbers of large dams (with reservoir area > 1 km^2) of different datasets. (a) Large dam

number in the GOODD dataset; (b) Large dam number in the GRanD dataset; (c) Large dam

number in the GeoDAR dataset; (d) Large dam number in the China-LDRL dataset; (e) Total

numbers of large dam from different datasets in China; (f) The relationships of large dam numbers

between China-LDRL, GRand and GeoDAR datasets.

658

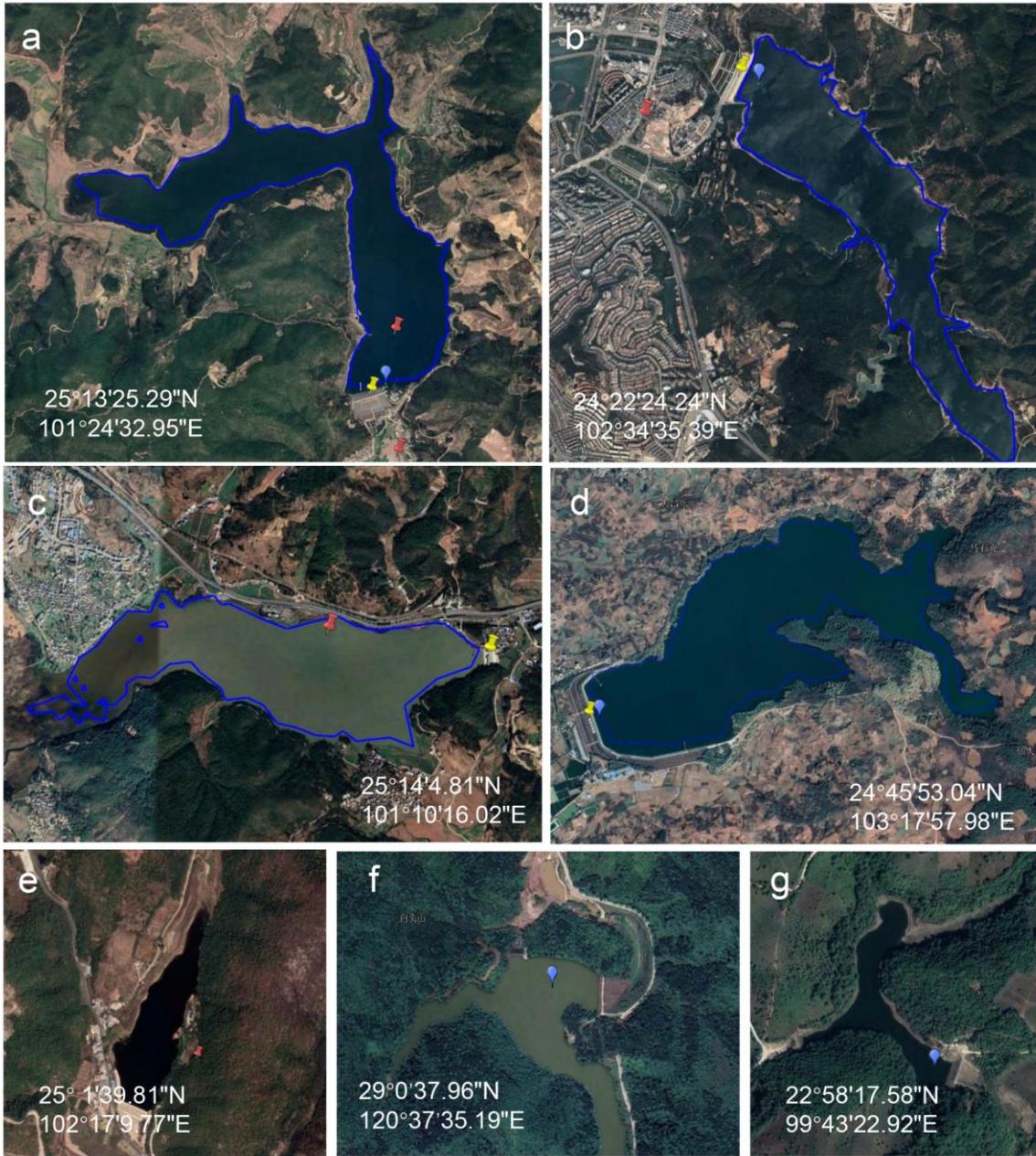



Fig. 10. Dams from the GOODD, GeoDAR, and China-LDRL datasets within Google Earth Pro

(© Google Earth Pro 2019).